THE BELL SYSTEM TECHNICAL JOURNAL
Vol. 57, No. 4, April 1978
Printed in U.S.A.

Loop Plant Modeling:

A Model of Cable Pairs Added at the Main Frames for a Large Entity

By W. N. BELL and S. BLUM

(Manuscript received August 17, 1977)

In estimating annual construction budgets and work forces required for a large entity, such as an operating company, it is desirable to know the total annual number of cable pairs to be added at the main frames in all of the feeder routes in the entity. These pairs are the result of many independent relief decisions made in the individual feeder routes. A model is presented which relates the total annual number of cable pairs added at the main frames to aggregate relief timing and sizing design parameters, an aggregate demand forecast and the total assigned and available pairs at the main frames in all of the routes in the entity. Because the model does not require individual route data or specific relief projects, it can be used as a "top-down" check on the "bottom-up" requirements determined by aggregating a list of identified relief projects. It is also useful for estimating cable pair requirements when specific relief projects are not yet identified.

I. INTRODUCTION

Growth in the number of telephone subscribers requires the periodic addition of cable pairs in the feeder network.[1] The feeder network comprises the large backbone cables which funnel cable pairs from the distribution network back toward the local wire center. The feeder cables are terminated in the wire center on a main distributing frame which serves as the interface between the loop network and the switching equipment.

At the main frame, the feeder cables are grouped into separate feeder routes which serve disjoint geographical areas within the boundaries of the wire center. Each route at the main frame is composed of a number

of available pairs, some of which are already assigned to existing subscribers served by the feeder route. The ratio of assigned pairs to available pairs at the main frame is called the main frame fill of the feeder route. The fill at other points along the route is defined in a similar manner.

When the fills in a portion of the feeder route become too high, new cable must be added. Most relief cables extend cable pairs already terminated on the main frame further out into the route. If tbe main frame fill of the route is also high, the new cable will be terminated on the main frame. This relief process is a complex one, involving both economic[2] and physical[3] considerations in the particular feeder route.

Individual relief timing[4] and sizing[2] decisions are made throughout the year by engineers based in local district engineering offices. A typical district has responsibility for about 100 main feeder routes. Districts are combined into administrative entities called areas which contain about 4 districts and 400 feeder routes. A typical operating company contains about 4 areas and 1600 feeder routes.

The aim of this paper is to present a model which has been derived to estimate the total number of cable pairs added on the main frames during a year in a large entity such as an area or company. The number of these pairs has historically been related to the total amount of feeder cable used in the entire feeder network during the year, and hence are an important component of an entity's annual construction budget and work force estimates. The model uses aggregate timing and sizing design parameters, an aggregate demand forecast and known aggregate main frame data such as total available pairs and total assigned pairs in the entity.

The model can be used to estimate future cable pair requirements for a large entity without identifying individual feeder routes or individual relief jobs. Hence the model is useful for estimating future cable pair requirements when specific relief projects are not yet identified, and for providing a simple "top-down" check on the "bottom-up" requirements obtained by aggregating a "market list" of identified relief projects. This check is useful because of the large number of feeder relief projects required each year in a large entity such as an area or company.

II. SUMMARY

A model is derived to estimate the total increase in available pairs at the main frames next year in all of the routes in a large entity. The model can be applied recursively to estimate the total increase in available pairs for any future year.

The problem of concern is formulated in Section III. Known aggregate main frame data are defined for an entity in terms of analogous data from the individual feeder routes. The basic approach used to derive the model

is outlined in Section 3.3 and involves modeling route relief at the main frame as a binomially distributed random variable where success connotes terminating a new cable on the main frame next year. Most of the modeling effort involves calculating the probability of relief (success) next year. Assumptions used to derive the model are conveniently summarized by defining an "analog entity" in Section 3.4.

The model is derived for an analog entity in three stages in Section IV. The initial result assumes that the fill at relief—the main frame fill in the route when a new cable is terminated—and cable size terminated at next relief will be the same as the fill at relief and cable size terminated at last relief in all of the routes in the analog entity. Adjustments to this steady state result are then derived in Sections 4.2 and 4.3 which enable the model to account for both changes in fill at relief and changes in cable size terminated.

The general model for estimating the increase in available pairs (ΔP) next year in an entity is

$$\overline{\Delta P} = \frac{G}{\hat{A}} + (1 - \theta)P\left(1 - \frac{F}{\hat{A}}\right) + N\Delta S$$

where

$$\hat{A} = \frac{W}{P - \frac{1}{2}NS}$$

is an estimate of the average fill at last relief for the entity. The estimates are expressed in terms of known aggregate main frame parameters for the entity which are either observed at the beginning of next year or forecast for next year:

N = number of routes
W = total assigned pairs
P = total available pairs
S = average cable size terminated
G = growth in assigned pairs next year
F = average fill at next relief
ΔS = change in average cable size terminated next year

The parameter θ depends on the impedance of the entity to a change in fill at relief.

Confidence intervals are calculated for the model in Section 4.5 as a function of the number of routes in the entity for the idealized case of identical routes. The resulting error bounds represent the best possible model performance in actual entities.

Finally, a validation study is described in Section V where the performance of the model is evaluated in actual entities. This study also

provided the means for empirically selecting a universal value of 0.6 for the impedance parameter θ. An error bound of 12 percent at the 90 percent confidence level was determined for the $\theta = 0.6$ model when applied to a typical company-sized entity of 1600 routes. The corresponding error bound at the 50 percent confidence level, or probable error, was 5 percent.

III. MODEL FORMULATION

A model for estimating the total increase in available pairs at the main frames in a given entity next year will be derived. This parameter is an indicator of the total feeder cable requirements in the entity next year. The goal of the modeling effort is to estimate this parameter using known aggregate main frame data and to determine the statistical accuracy of the estimate when used in actual entities.

Known aggregate main frame data for an entity include:

- number of feeder routes
- total assigned pairs
- total available pairs
- average cable size terminated
- growth forecast
- average fill at next relief
- change in average cable size terminated

These data are either aggregated annually from the feeder routes which compose the entity, or are forecast directly for the entity. The above parameters will be formally defined in terms of analogous route parameters in Section 3.2.

3.1 Basic route parameters

A feeder route is described at the main frame by the following set of basic route parameters (denoted by lower-case letters) which are either observed or forecast at the beginning of next year (see Fig. 1):

$$\left.\begin{array}{l} w = \text{assigned pairs} \\ p = \text{available pairs} \\ s = \text{cable size terminated} \\ a = \text{fill at last relief} \end{array}\right\} \quad \begin{array}{l}\text{observed at the} \\ \text{beginning of next year}\end{array}$$

g = growth in assigned pairs next year
f = fill at next relief

The ratio w/p is the main frame fill at the beginning of next year. The last cable terminated at the main frame contained s pairs and was placed in service when the main frame fill, or fill at last relief, was a. A new cable will be terminated in the route when the main frame fill reaches the fill at next relief, f. Initially, we will assume that the new cable contains the same number of pairs as the last cable terminated (s). This assumption will be relaxed in Section 4.3.

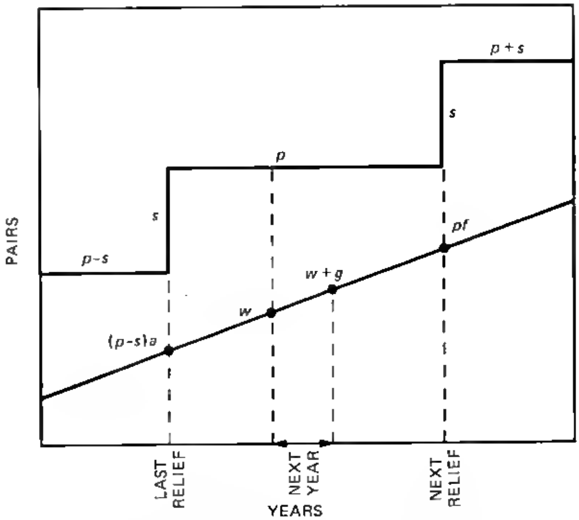

Fig. 1—Feeder route parameters.

3.2 Known aggregate main frame data

Known aggregate main frame data (denoted by upper-case letters) compiled from the individual feeder routes at the beginning of next year are

N = number of feeder routes
$W = \sum w_j$ = total assigned pairs
$P = \sum p_j$ = total available pairs

$$S = \frac{1}{N} \sum s_j = \text{average cable size terminated}$$

The parameters W and P are simply aggregate facility data which have historically been collected to monitor the usage of the feeder network at the main frames. The ratio W/P is the main frame fill of the entity at the beginning of next year. The average cable size terminated (S) is an aggregate sizing design parameter.

The following parameters are forecast for the entity for next year:

$G = \sum g_j$ = growth in assigned pairs next year

$$F = \frac{1}{P} \sum p_j f_j = \text{average fill at next relief.}$$

G is a standard demand forecast and F is an aggregate timing design parameter. These parameters are typically forecast directly for the entity rather than aggregated from the forecasts of the analogous parameters in each route.

The definition of N, W, P and G are intuitively aggregates of the analogous feeder route parameters. The definition of the aggregate de-

sign parameters S and F will become clear during the derivation of the model (Section IV). Note that there is no known aggregate parameter which is analogous to the fill at last relief (a) in a single route (see Section 3.1). This unknown aggregate parameter will be introduced with the definition of the analog entity in Section 3.4.

The basic approach used to derive a model for estimating the total increase in available pairs next year in an entity is described in the next section.

3.3 Basic approach

The increase in available pairs (Δp) next year in a route is either s or zero, depending on whether a new cable is terminated next year or not. Hence Δp can be expressed as χs where χ is the relief indicator for the route next year, defined as

$$\chi = \begin{cases} 1 & \text{if a new cable is} \\ & \text{terminated next year} \\ 0 & \text{otherwise} \end{cases}$$

The increase in available pairs next year in an entity can be expressed as

$$\begin{aligned} \Delta P &= \sum \Delta p_j \\ &= \sum \chi_j s_j \end{aligned} \tag{1}$$

The minimum value of $\Delta P = 0$ corresponds to the unlikely case of no new cable being terminated next year ($\chi_j = 0 : j = 1,N$), while the maximum value of $\Delta P = NS$ corresponds to the equally unlikely case of a new cable being terminated in every route next year ($\chi_j = 1 : j = 1,N$).

The relief indicator (χ) for a route next year can be modeled as a binomially distributed random variable where success corresponds to a new cable being terminated next year ($\chi = 1$). Treating the cable size (s_j) as a known parameter, the expected increase in available pairs next year in the entity can be expressed as

$$\overline{\Delta P} = \sum \lambda_j s_j \tag{2}$$

where $\lambda_j = \overline{\chi}_j$ is the probability of relief next year for route j.

Intuitively, the probability of relief next year should depend on how far the route is through its relief cycle at the beginning of next year. The fraction through the relief cycle (r) decomposes the relief cycle (t) into two components; the time since last relief (rt) and the time to next relief $[(1 - r)t]$. The route will be relieved next year if the time to next relief (observed at the beginning of next year) is one year or less. If r is modeled

as a random variable, then the probability of relief next year,

$$\lambda = \mathcal{P}\Big((1-r)t \le 1\Big)$$
$$= \mathcal{P}\left(r \ge 1 - \frac{1}{t}\right) \tag{3}$$

depends on the tail of the distribution of the fraction through the relief cycle at the beginning of next year.

The route is said to be in the steady state if the fill at next relief is the same as the fill at last relief ($f = a$). In this case (see Fig. 2), the relief cycle of the route—the number of years between last and next relief—is

$$t^{ss} = \frac{sa}{g} \tag{4}$$

and the fraction through the relief cycle at the beginning of next year is

$$r^{ss} = \frac{w - (p-s)a}{sa} \tag{5}$$

If the fill at next relief (f) forecast at the beginning of next year is not equal to the fill at last relief (a), then the relief cycle becomes

$$t = \frac{p(f-a) + sa}{g} \tag{6}$$

If $f > a$ (see Fig. 3), then the relief cycle is increased and the fraction

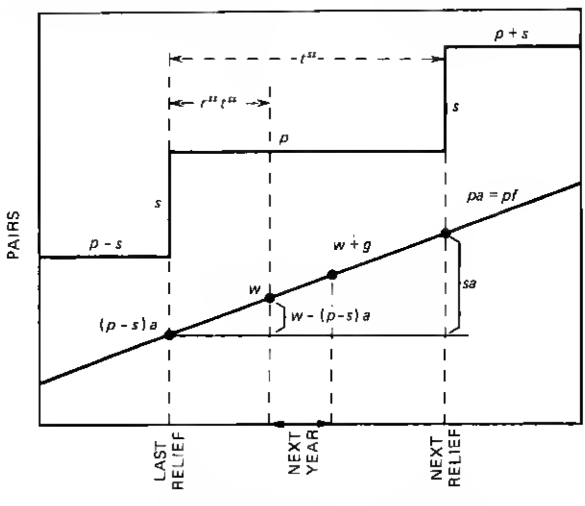

Fig. 2—Feeder route in steady state ($f = a$).

through the relief cycle is decreased such that their product, rt = time since last relief, remains constant. Because the time since last relief is also equal to $r^{ss}t^{ss}$, we have $rt = r^{ss}t^{ss}$ or

$$r = r^{ss}\frac{t^{ss}}{t} \tag{7}$$

If r^{ss} is modeled as a random variable, then the probability of relief next year is

$$\begin{aligned} \lambda &= \mathcal{P}\left(r \geq 1 - \frac{1}{t}\right) \\ &= \mathcal{P}\left(r^{ss} \geq \frac{t}{t^{ss}} - \frac{1}{t^{ss}}\right) \end{aligned} \tag{8}$$

Hence r^{ss}—the fraction through the relief cycle *if* the route is in the steady state ($f = a$)—is the basic random variable which will be used to derive the model.

The distribution of r^{ss} determines the distribution of w—the assigned pairs in the route at the beginning of next year—which can be expressed in terms of r^{ss} by solving eq. (5) for w,

$$w = ap + as(1 - r^{ss}) \tag{9}$$

This is the only basic route parameter defined in Section 3.1 which is modeled as a random variable. All other parameters (p,s,a,g,f) are as-

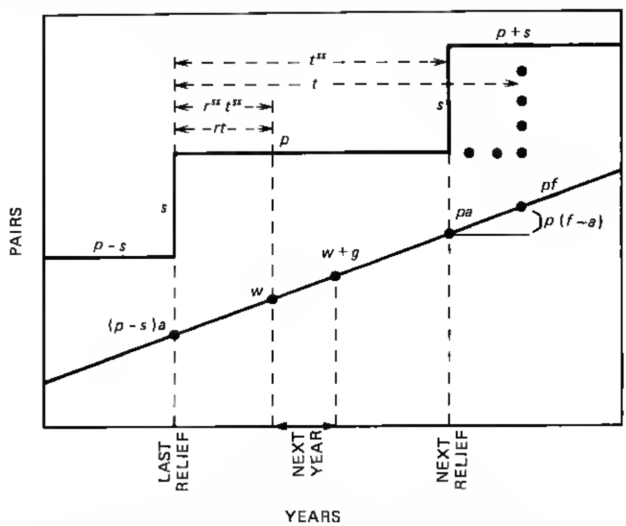

Fig. 3—Feeder route with a change in fill at relief ($f > a$).

sumed to be deterministic. Hence the length of the relief cycle [t^{ss} and t, see eqs. (4) and (6)] are deterministic, but the fraction through the relief cycle [r^{ss} and r, see eqs. (5) and (7)] are random variables.

3.4 Model assumptions—the analog entity

Several assumptions are required to derive a model for the total increase in available pairs (ΔP) next year in an entity in terms of the known aggregate main frame data (N,W,P,S,G,F) defined in Section 3.2. These assumptions are needed because the distribution of the aggregate data over the routes in the entity ($w_j, p_j, s_j, a_j, g_j, f_j : j = 1,N$) is in general not known. The assumptions are:

A1: All of the routes in the entity are assumed to have the same fill at last relief ($a_j = A : j = 1,N$) where A is an unknown parameter that must be estimated in terms of known aggregate data.

A2: If the routes in the entity are in the steady state condition ($f_j = A : j = 1,N$), then the fractions through the relief cycle of the routes ($r_j^{ss} : j = 1,N$) are independent and identically distributed uniform random variables on the unit interval. In other words, if a route is selected at random from routes in the steady state, then we are equally likely to observe any value of r^{ss} between 0 (just after relief) and 1 (just before relief).

An entity satisfying these two conditions will be referred to as an analog entity. The model will be derived for an analog entity in the next section. The performance of this model in actual entities will be studied in Section V.

IV. AVAILABLE PAIR INCREASE MODEL

A model for estimating the increase in available pairs next year in an analog entity is derived. This model can be applied recursively to obtain an estimate of the increase in available pairs for any future year.

4.1 Steady state model

The expected increase in available pairs in the analog entity next year [eq. (2)] is

$$\overline{\Delta P} = \sum \lambda_j s_j$$

where s_j is the cable size terminated and λ_j is the probability of relief next year in route j. If the routes in the analog entity are in the steady state condition (i.e., $f_j = A : j = 1,N$), then $t_j = t_j^{ss}$ [see eqs. (4) and (6)] and the probability of relief next year [eq. (8)] is

$$\begin{aligned}\lambda_j &= \mathcal{P}\left(r_j^{ss} \geq \frac{t_j}{t_j^{ss}} - \frac{1}{t_j^{ss}}\right)\\ &= \mathcal{P}\left(r_j^{ss} \geq 1 - \frac{1}{t_j^{ss}}\right)\end{aligned}$$

Since r_j^{ss} is assumed to be uniformly distributed on the unit interval (assumption A2), we have

$$\lambda_j = \begin{cases} 1 & \text{if} \quad t_j^{ss} < 1 \\ \dfrac{1}{t_j^{ss}} & \text{if} \quad t_j^{ss} \geq 1 \end{cases} \tag{10}$$

The physical significance of $t_j^{ss} < 1$ is that the last cable terminated (s_j) was sized so small that the route requires another relief cable in less than one year. This is not a normal engineering practice and it will not be considered in the model.

The expected available pair increase in route j for the case $t_j^{ss} \geq 1$ is

$$\begin{aligned} \overline{\Delta p_j} &= \lambda_j s_j \\ &= \frac{s_j}{t_j^{ss}} \\ &= \frac{g_j}{A} \end{aligned} \tag{11}$$

This result follows from eqs. (10) and (4) and assumption A1. Hence if s_j pairs are terminated every t_j^{ss} years, then s_j/t_j^{ss} pairs will be terminated on average next year. Equivalently, if cables were sized for only one year's growth, then g_j/A pairs would be terminated next year.

The expected increase in available pairs in the analog entity next year is obtained by summing eq. (11) over all routes,

$$\begin{aligned} \overline{\Delta P} &= \sum \frac{g_j}{A} \\ &= \frac{G}{A} \end{aligned} \tag{12}$$

where G is the entity growth forecast for next year. In the following sections, we derive adjustments to this simple steady state model by relaxing some of the restrictive assumptions leading to this result.

4.2 Change in fill at relief

The steady state assumption will now be relaxed by allowing the forecast fill at next relief in route j (f_j) to be different from the fill at last relief (A). Such a change could take place if locally selected fills at relief were used in place of a constant fill at relief policy represented by A.

4.2.1 Instantaneous change in fill at relief

If the fill at next relief is changed from A to f_j at the beginning of next year, then the probability of relief next year in route j [eq. (8)] is

$$\lambda_j = \mathcal{P}\left(r_j^{ss} \geq \frac{t_j}{t_j^{ss}} - \frac{1}{t_j^{ss}}\right)$$

$$= \begin{cases} 0 & \text{if} \quad t_j > t_j^{ss} + 1 \\ \dfrac{1}{t_j^{ss}} + 1 - \dfrac{t_j}{t_j^{ss}} & \text{if} \quad 1 \leq t_j \leq t_j^{ss} + 1 \\ 1 & \text{if} \quad t_j < 1 \end{cases} \tag{13}$$

This result follows because r_j^{ss} is uniformly distributed on the unit interval by assumption A2. If $f_j \gg A$ such that a year or more is added to the relief cycle (case $t_j > t_j^{ss} + 1$), then the route will not be relieved next year ($\lambda_j = 0$). On the other hand, if $f_j \ll A$ such that the relief cycle is reduced to less than one year (case $t_j < 1$), then the route will be relieved next year ($\lambda_j = 1$).

The expected increase in available pairs in route j for the case $1 \leq t_j \leq t_j^{ss} + 1$ is

$$\begin{aligned} \overline{\Delta p_j} &= \lambda_j s_j \\ &= \left(\frac{1}{t_j^{ss}} + 1 - \frac{t_j}{t_j^{ss}}\right) s_j \\ &= \frac{g_j}{A} + p_j\left(1 - \frac{f_j}{A}\right) \end{aligned} \tag{14}$$

This result follows after substituting eq. (13) for λ_j and eqs. (4) and (6) for t_j^{ss} and t_j. Note that $\overline{\Delta p_j}$ is greater than the steady state value (g_j/A) if $f_j < A$ and is less than the steady state value if $f_j > A$.

The expected increase in available pairs next year in the analog entity is obtained by summing eq. (14) over all routes,

$$\begin{aligned} \overline{\Delta P} &= \sum\left(\frac{g_j}{A} + p_j\left(1 - \frac{f_j}{A}\right)\right) \\ &= \frac{G}{A} + P\left(1 - \frac{1}{P}\frac{\sum p_j f_j}{A}\right) \\ &= \frac{G}{A} + P\left(1 - \frac{F}{A}\right) \end{aligned} \tag{15}$$

where F is the average fill at next relief for the original entity. The previous steady state result [G/A, eq. (12)] is adjusted by a factor that depends on the change in fill at relief. The adjustment is positive if $F < A$ and negative if $F > A$.

4.2.2 Impedance to a change in fill at relief

The last result [eq. (15)] was derived by assuming an instantaneous change in fill at relief from A to f_j in all routes in the analog entity at the beginning of next year. In practice, we should not expect an instantaneous change. It is not practical or economical to defer a job (by increasing the fill at next relief) which is already under construction or for which material has already been ordered. On the other hand, it is not possible to suddenly advance a job (by decreasing the fill at next relief) that had been planned for further into the future because major feeder relief jobs must be engineered a year or more in advance to allow time for job approval, ordering material and construction.

These considerations can be accounted for in the model by introducing the concept of the impedance of the analog entity to a change in fill at relief. The effect of this impedance is to cause the fill at relief to be changed gradually rather than instantaneously in all routes. In this case, the average fill at next relief (F') for the analog entity is somewhere between A and F—the average fill at next relief for the original entity. Hence F' can be expressed as

$$F' = \theta A + (1 - \theta)F \tag{16}$$

for some θ between 0 and 1. The parameter θ is related to the impedance of the analog entity to the change in fill at relief. If $\theta = 0$, the analog entity offers zero impedance and the change is instituted instantaneously in all routes ($F' = F$). If $\theta = 1$, then the analog entity offers infinite impedance and the change is never instituted ($F' = A$).

The expected available pair increase in the analog entity next year based on F' rather than F is

$$\begin{aligned}\overline{\Delta P}(\theta) &= \frac{G}{A} + P\left(1 - \frac{F'}{A}\right) \\ &= \frac{G}{A} + (1 - \theta)P\left(1 - \frac{F}{A}\right)\end{aligned} \tag{17}$$

Note that θ attenuates the effect of the change in fill at relief on the expected increase in available pairs. Equation (17) reduces to our previous result [eq. (15)] if the fill at relief is changed instantaneously in all routes ($\theta = 0$). An empirical value of θ will be determined in Section V based on the observed behavior of actual entities to a change in fill at relief.

4.3 Change in average cable size terminated

In all of the results presented so far, we have assumed that the same sized cable is terminated whenever a particular feeder route is relieved. In practice (see Ref. 2), changes in many factors can cause future cable sizes to be different from those placed in the past. In this case, the new cables terminated next year would cause a change in average cable size

terminated from its value at the beginning of the year (S). We will now derive another adjustment to the model by determining the impact of a change in average cable size (ΔS) on the expected increase in available pairs next year.

The change in average cable size terminated next year can be defined in the original entity as

$$\Delta S = \frac{1}{N} \sum \chi_j (c_j - s_j) \tag{18}$$

where N is the number of routes in the entity, χ_j is the relief indicator function for next year, s_j is the cable size terminated at last relief and c_j is the cable size to be terminated at next relief in route j. We will assume that a change in average cable size next year is forecast directly for the entity.

If the cable sizes ($c_j : j = 1,N$) are carried over to the routes of the analog entity, then the expected increase in available pairs next year can be expressed as

$$\begin{aligned}\Delta P &= \sum \chi_j c_j \\ &= \sum \chi_j (s_j + c_j - s_j) \\ &= \sum \chi_j s_j + \sum \chi_j (c_j - s_j) \\ &= \sum \chi_j s_j + N \Delta S\end{aligned}$$

The first term is the increase in available pairs in the analog entity if there is no change in average cable size next year. The model for its expected value is given in eq. (17). If average cable size is forecast to change by ΔS next year, then the increase in available pairs will change by $N\Delta S$.

The general model is obtained by adding $N\Delta S$ to eq. (17):

$$\overline{\Delta P}(\theta) = \frac{G}{A} + (1 - \theta)P\left(1 - \frac{F}{A}\right) + N\Delta S \tag{19}$$

This result represents the final adjustment to the original steady state model of G/A. The general model can be applied recursively to obtain the expected increase in available pairs for any future year.

4.4 Fill at last relief model

When defining the analog entity, we assumed that the fill at last relief was the same in all routes. A value for this parameter (A) is needed to calculate the expected increase in available pairs next year in the analog entity [see eq. (19)]. In this section, we derive an estimate of A in terms of known aggregate main frame data (see Section 3.2).

Because A is the fill at *last* relief in the analog entity, its estimator can only depend on parameters which can be observed at the beginning of

next year. These include the basic route parameters $(w_j, p_j, s_j : j = 1,N)$ and the corresponding aggregate parameters (N,W,P,S).

The basic relationship in the analog entity between the route parameters (w_j, p_j, s_j) and A is expressed in eq. (9),

$$w_j = Ap_j - As_j(1 - r_j^{ss}) \tag{20}$$

Both r_j^{ss} and w_j are modeled as random variables and r_j^{ss} is assumed to be uniformly distributed on the unit interval.

Summing eq. (20) over all routes in the analog entity yields

$$N\overline{w} = AP - ANS(1 - R^{ss})$$

which can be solved for A to obtain

$$A = \frac{N\overline{w}}{P - NS(1 - R^{ss})} \tag{21}$$

In the last two equations, $\overline{w} = (\Sigma w_j)/N$ and

$$R^{ss} = \frac{1}{N} \sum (s_j/S) r_j^{ss} \tag{22}$$

are random variables and N, P and S are observed aggregate parameters at the beginning of next year. If $\overline{w}$ is observed to be W/N at the beginning of next year, then a random variable (α) can be defined as an estimator of A where

$$\alpha = \frac{W}{P - NS(1 - R^{ss})} \tag{23}$$

An estimate of A is then the expected value of α.

The expected value of α does not have a simple form because the random variable R^{ss}—which is analogous to the fraction through the relief cycle in a single route—is approximately normally distributed. This follows from a generalized version of the Central Limit Theorem, called Lindebergs Theorem,[5] for uniformly bounded independent random variables. In our case, R^{ss} is the sample mean of the independent random variables $[(s_j/S)r_j^{ss} : j = 1,N]$ which are uniformly bounded by $k = s_{max}/S$ where s_{max} is the largest manufactured cable size. The expected value and variance of R^{ss} can be calculated from eq. (22) and the expected value ($\frac{1}{2}$) and variance ($\frac{1}{12}$) of r_j^{ss},

$$\begin{aligned} \mathcal{E}(R^{ss}) &= \frac{1}{N} \sum (s_j/S)\mathcal{E}(r_j^{ss}) \\ &= \frac{1}{2} \\ \text{var}(R^{ss}) &= \frac{1}{N^2} \sum (s_j/S)^2 \text{var}(r_j^{ss}) \\ &= \frac{1}{12N^2} \sum (s_j/S)^2 \end{aligned} \tag{24}$$

An estimate of A, $\hat{A}$, can be obtained by replacing R^{ss} in eq. (23) by its expected value of ½, yielding

$$\hat{A} = \frac{W}{P - \frac{1}{2}NS} \tag{25}$$

This is a reasonable estimator for A if the variance of R^{ss} is small. A bound on var(R^{ss}) can be derived from the bound on s_j/S. Recall that $s_i/S \le k$, which together with eq. (24) yields

$$\text{var}(R^{ss}) \le \frac{k^2}{12N}$$

Using a value of $k = s_{\max}/S = 4$ which corresponds to a typical average cable size terminated $S = 900$ pairs and the largest manufactured cable size $s_{\max} = 3600$ pairs yields var(R^{ss}) ≤ 0.003 for a typical area sized entity of 400 routes. Hence for large entities, $\hat{A}$ [eq. (25)] is a reasonable estimator of A.

This estimator will be used for A in the general model [eq. (19)] for estimating the available pair increase next year in the analog entity. A confidence interval for the model will be derived in the next section for the idealized case of an entity of identical routes. The performance of the model in actual entities will be analyzed in Section V.

4.5 Idealized model accuracy

A confidence interval for the expected increase in available pairs next year will be derived for an entity of N identical routes where the probability of relief next year and the cable size terminated are the same in all routes. The calculated error bounds for this idealized case are representative of the best possible model performance in actual entities and will be compared to observed model errors in Section V.

If the probability of relief next year and the cable size terminated are the same in all routes, then the increase in available pairs next year can be represented as a sum of N independent and identically distributed binomial random variables,

$$\Delta P = \sum \chi_j s$$

The expected value and variance of ΔP are

$$\mathcal{E}(\Delta P) = N\lambda s \tag{26}$$

$$\text{var}(\Delta P) = Ns^2\lambda(1 - \lambda) \tag{27}$$

where λ is the probability of relief next year and s is the cable size ter-

minated in all routes. If N is large, ΔP is approximately normally distributed by the De Moivre-Laplace Theorem.[5]

A δ-confidence interval for the relative deviation of ΔP from its expected value can be obtained from the probability statement,

$$\mathcal{P}\left(\left|\frac{\Delta P - \mathcal{E}(\Delta P)}{\mathcal{E}(\Delta P)}\right| < \epsilon\right) \geq \delta$$

which can be expressed in terms of a unit normal random variable,

$$\mathcal{P}\left(\left|\frac{\Delta P - \mathcal{E}(\Delta P)}{(\mathrm{var}(\Delta P))^{1/2}}\right| < \epsilon \frac{\mathcal{E}(\Delta P)}{(\mathrm{var}(\Delta P))^{1/2}}\right) \geq \delta$$

If β is selected such that

$$\mathcal{P}\left(\left|\frac{\Delta P - \mathcal{E}(\Delta P)}{(\mathrm{var}(\Delta P))^{1/2}}\right| < \beta\right) \geq \delta$$

then the error bound ϵ can be expressed as

$$\epsilon = \beta \frac{(\mathrm{var}(\Delta P))^{1/2}}{\mathcal{E}(\Delta P)}$$

Substituting eqs. (26) and (27) for $\mathcal{E}(\Delta P)$ and $\mathrm{var}(\Delta P)$ yields

$$\epsilon(N,\lambda) = \beta \left(\frac{1-\lambda}{\lambda N}\right)^{1/2} \tag{28}$$

This result will be used in the next section as a scale factor to compare observed errors in entities with different numbers of routes and average probabilities of relief.

Error bounds $\epsilon(N,\lambda)$ at the 50 percent ($\delta = 0.5, \beta = 0.675$) and 90 percent ($\delta = 0.9, \beta = 1.645$) confidence levels for an entity of identical routes with a probability of relief next year of $\lambda = 0.2$ are given in Table I as a function of the number of routes in the entity. These results are representative of the best possible model performance in actual entities.

V. VALIDATION STUDY

A validation study was performed to determine the statistical accuracy of the model when applied in actual entities. This analysis will also provide the means of calibrating the model by empirically selecting a

Table I — Idealized error bounds for available pair increase model ($\lambda = 0.2$)

Number of routes	Confidence level 50%	Confidence level 90%
100	14%	33%
200	10	23
400	7	16
1600	3	8

value for θ based on the observed behavior of actual entities to changes in fill at relief.

5.1 Validation plan

Route facility charts provide the data required to perform the validation study. These charts are maintained for each route in some areas to record the past history of assigned and available pairs and to project future growth and relief requirements at tbe main frame. A typical route facility chart is shown in Fig. 4. The assigned and available pairs in the route at the main frame are counted at year-end and posted on the facility chart. An increase in available pairs in the route caused by terminating a new cable on the main frame is shown in the month that the pairs were made available for service. In this case, an estimate of the fill at relief can be made by linearly interpolating between the year-end counts of assigned pairs made before and after the service date of the job.

The validation study consisted of applying the model in a number of sample entities and then analyzing the statistics of the model errors. The aggregate main frame data $(N, W, P, S, G, F, \Delta S)$ required by the model can be compiled from facility chart data on each of the routes in a sample entity. The equations in Sections 3.2 and 4.3 were used with one exception: The average fill at next relief (F) was calculated using only routes that were relieved next year. Since historical data were used in the study,

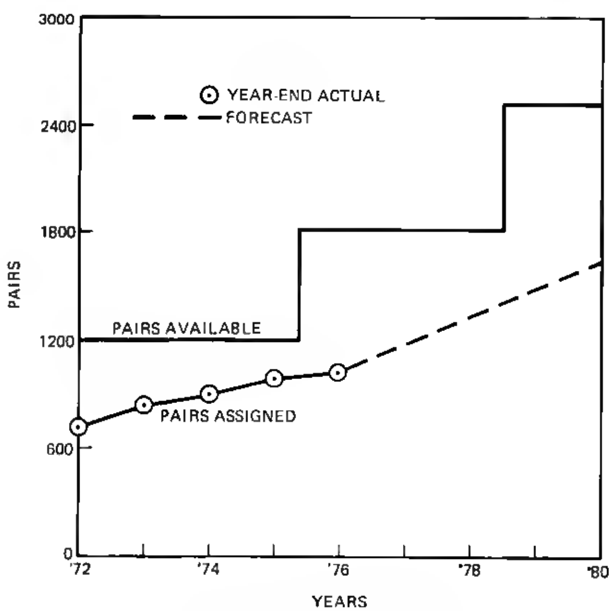

Fig. 4—Route facility chart.

actual values could be used for G, ΔS and F rather than forecast values which had the desirable effect of eliminating forecast errors from the validation study.

The general model [eq. (19)] was then used to calculate an estimate of the increase in available pairs next year in the sample entity for various values of θ between zero and one. An estimated value ($\hat{A}$) of the fill at last relief (A) required by the model was calculated using eq. (25). The actual model error for each value of θ,

$$e(\theta) = \left| \frac{\overline{\Delta P}(\theta) - \Delta P}{\overline{\Delta P}(\theta)} \right|$$

was then scaled using the equation

$$\tilde{e}(\theta) = e(\theta) \frac{\epsilon(200{,}0.2)}{\epsilon(N,\lambda)}$$

where $\epsilon(N,\lambda)$ is given by eq. (28) and λ is the fraction of routes in tbe sample entity which were actually relieved next year. The scaling is intended to make errors comparable in sample entities with different numbers of routes (N) and probabilities of relief (λ). The scale factor is arbitrarily defined to be unity when $N = 200$ and $\lambda = 0.2$.

The distribution of scaled model errors over the sample entities was determined for each value of θ. The behavior of the 90 percent point and the 50 percent point, or median, of these distributions as a function of θ was used as the basis for selecting the best value for θ. The 50 percent point of the distribution of signed scaled errors was also studied as a function of θ to help detect model biases. Similar studies were performed on subsets of the sample entities to determine if the model could be calibrated using a universal value of θ.

5.2 Validation study results

Facility chart history on available and assigned pairs from over 650 routes were gathered from three different areas (see Table II). History was sought from year-end 1968 to year-end 1975 but this amount of history was not available in all routes.

Sample entities were defined within an area for each year by selecting routes whose growth in assigned pairs during the year was at least one percent of the total assigned pairs at the beginning of the year. Tbe validation study was restricted to only growth routes because tbe con-

Table II — Route data

Area	Number of routes
A	283
B	211
C	168

cepts of timing and sizing relief are most meaningful in a growing route.

This process yielded 21 sample entities (i.e., 7 different years in 3 different areas). The actual number of routes in each sample entity are shown in Table III. Model errors were calculated as described in the validation plan (Section 5.1). The 90 and 50 percent points of the distribution of scaled errors are plotted versus θ in Fig. 5. The 50 percent point of the distribution of signed scaled errors—which ideally should be zero—is also shown to help detect model biases.

The best value of θ for estimating the increase in available pairs is 0.6 which minimizes the 90 percent point of the distribution of scaled errors. At this value of θ, the 50 percent point of the distribution is nearly minimized, and the 50 percent point of the distribution of signed scaled errors is almost zero. Similar analyzes performed for the 7 sample entities in each area yielded values of θ between 0.5 and 0.7. These results indicate that it is possible to calibrate the model with a universal value of $\theta = 0.6$.

The anticipated accuracy of the best model ($\theta = 0.6$) when applied in

Table III — Number of routes in sample entities

Area	Year '69	'70	'71	'72	'73	'74	'75
A	155	166	184	197	197	184	219
B	60	81	98	128	171	166	166
C	42	50	75	81	92	88	125

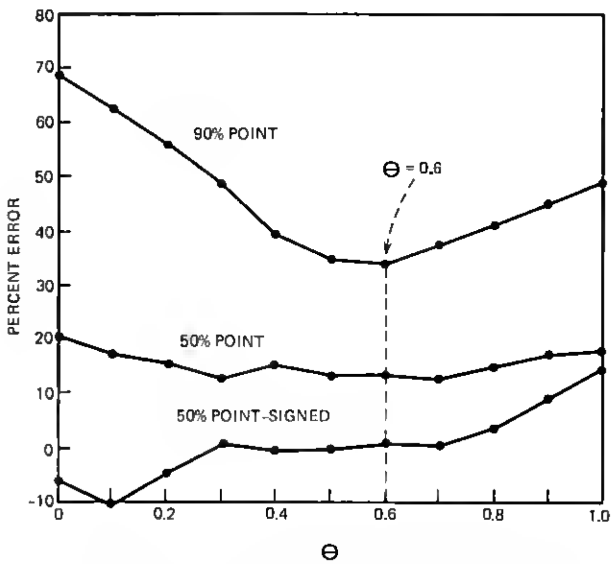

Fig. 5—Scaled model errors ($N = 200$, $\lambda = 0.2$).

Table IV — Anticipated error bounds for the available pair increase model ($\theta = 0.6$, $\lambda = 0.2$)

Number of routes	Confidence level 50%	90%
100	19%	48%
200	13	34
400	10	24
1600	5	12

an entity of N routes with an average probability of relief of λ can be expressed in terms of the scaled errors as

$$e(N,\lambda) = \bar{e}(\theta = 0.6)\,\frac{\epsilon(N,\lambda)}{\epsilon(200,0.2)}$$

where $\epsilon(N,\lambda)$ is given in eq. (28) and $\bar{e}(\theta = 0.6)$ is read from Fig. 5 for either the 50 percent or 90 percent confidence level. Anticipated error bounds as a function of the number of routes in entities with a probability of relief of 0.2 are summarized in Table IV. These results are somewhat higher than the corresponding error bounds which were calculated for an entity of identical routes (see Table I) because of route-to-route variations in the sample entities.

VI. ACKNOWLEDGMENTS

We are grateful to R. L. Young for his comments on a preliminary draft of this paper. L. G. McMillan, A. E. Gibson, and T. P. Shpiz developed the data analysis techniques and computer programs required for the validation study.

REFERENCES

1. N. G. Long, "Loop Plant Modeling: Overview," B.S.T.J., this issue.
2. J. Freidenfelds, "A Simple Model for Studying Feeder Capacity Expansion," B.S.T.J., this issue.
3. B. L. Marsh, "The Feeder Allocation Process," B.S.T.J., this issue.
4. W. L. G. Koontz, "An Approach to Modeling Operating Costs in the Loop Network," B.S.T.J., this issue.
5. W. Feller, *An Introduction to Probability Theory and Its Applications*, Vol. 1, New York: John Wiley and Sons, Inc., 1957, p. 172 and p. 239.